DOCUMENT RESUME

ED 197 308 CS 005 842

TITLE Testing and Evaluation in Reading and Communication Skills: Abstracts of Doctoral Dissertations Published in "Dissertation Abstracts International," July through December 1980 (Vol. 41 Nos. 1 through 6).

INSTITUTION ERIC Clearinghouse on Reading and Communication Skills, Urbana, Ill.

PUB DATE 80

NOTE 18p.: Pages may be marginally legible.

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Annotated Bibliographies: Cloze Procedure; Communication Skills; *Doctoral Dissertations; Elementary Secondary Education; Higher Education; Informal Reading Inventories; Interaction; *Reading Comprehension; *Reading Diagnosis; Reading Instruction; *Reading Research; Reading Skills; *Reading Tests; Standardized Tests; *Testing

ABSTRACT

This collection of abstracts is part of a continuing series providing information on recent doctoral dissertations. The 27 titles deal with a variety of topics, including the following: (1) electrodermal responses and informal reading criteria; (2) a comparison of standardized reading test scores, informal reading inventory results, and miscue analysis for above and below average third grade students; (3) the relationship between the cloze test and a hierarchical model of reading comprehension skills; (4) the relationship of reading achievement to certain selected tests and newly constructed measures of psychological processes; (5) the utility of selected interaction analysis systems as feedback tools during reading instruction; (6) an ethnographic instrument to assess the linguistic interactive behavior of classroom groups; (7) determining what is measured by multiple choice tests of reading comprehension; (8) the development and evaluation of a words-in-context test to determine instructional reading levels; (9) a canonical correlational analysis of the reading-writing relationship; (10) the effects of differing response criteria on the assessment of writing competence; and (11) oral reading errors of fifth grade students. (FL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED197308

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Testing and Evaluation
in
Reading and Communication Skills:

Abstracts of Doctoral Dissertations Published in Dissertation Abstracts International, July through December 1980 (Vol. 41 Nos. 1 through 6)

Compiled by the staff of the

ERIC Clearinghouse on Reading and Communication Skills

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
University Microfilms
International

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS 505 842

The dissertation titles contained here are published with permission of the University Microfilms International, publishers of Dissertation Abstracts International (copyright © 1980 by University Microfilms International), and may not be reproduced without their prior permission.

This bibliography has been compiled as part of a continuing series designed to make information on relevant dissertations available to users of the ERIC system. Monthly issues of Dissertation Abstracts International are reviewed in order to compile abstracts of dissertations on related topics, which thus become accessible in searches of the ERIC data base. Ordering information for the dissertations themselves is included at the end of the bibliography.

Abstracts of the following dissertations are included in this collection:

Al-Dahiry, Saleem Abdul Abbas Yasin
COMPARISON OF ORAL READING, SILENT READING, AND LISTENING COMPREHENSION OF THIRD GRADE STUDENTS ASSESSED BY GATES-MACGINITIE READING TEST

Bergman, Irwin B.
THE EFFECTS OF TEST-TAKING INSTRUCTION VS. PRACTICE WITHOUT INSTRUCTION ON THE EXAMINATION RESPONSES OF COLLEGE FRESHMEN TO MULTIPLE-CHOICE, OPEN-ENDED, AND CLOZE TYPE QUESTIONS

Bonds, Charles Wesley
ELECTRODERMAL RESPONSES AND INFORMAL READING CRITERIA

Brinzer, Raymond Joseph
EFFECTS OF A RATE EMPHASIS ON THE READING PERFORMANCE OF REFLECTIVE AND IMPULSIVE SIXTH-GRADE MALES AND FEMALES

Button, Linda Jean Orvis
STANDARDIZED READING TEST SCORES AND INFORMAL READING INVENTORY RESULTS WITH MISCUE ANALYSIS FOR ABOVE AND BELOW AVERAGE THIRD GRADE STUDENTS: A COMPARISON

Calica, Diana Krakower
A STUDY OF THE RELATIONSHIP BETWEEN THE CLOZE TEST AND A HIERARCHICAL MODEL OF READING COMPREHENSION SKILLS

Clement, Betty Ann Waidlich
A COMPARISON OF A STANDARDIZED GROUP READING TEST AND AN INFORMAL READING INVENTORY IN ESTIMATING AN INSTRUCTIONAL READING LEVEL

Davidson, Emma Sue
AN ANALYSIS OF FOURTH AND SIXTH GRADE READER PERFORMANCE USING CLOZE TESTS, GROUP READING INVENTORIES, AN INFORMAL READING INVENTORY, AND A STANDARDIZED READING TEST

DeSoto, Janet Louise
RELATIONSHIP OF READING ACHIEVEMENT TO CERTAIN SELECTED TESTS AND NEWLY CONSTRUCTED MEASURES OF PSYCHOLOGICAL PROCESSES

Edwards, Kathryn Inez
THE UTILITY OF SELECTED INTERACTION ANALYSIS SYSTEMS AS FEEDBACK TOOLS DURING THE READING INSTRUCTION PERIOD

Forest, Robert Gerard
AN ETHNOGRAPHIC INSTRUMENT TO ASSESS THE LINGUISTIC INTERACTIVE BEHAVIOR OF CLASSROOM GROUPS

Grant, Norma Louise
THE EFFECTS OF PASSAGE LENGTH AND READING MODE OF RETELLINGS AND QUESTION RESPONSES OF SIXTH GRADE CHILDREN

Haertel, Edward Henry
DETERMINING WHAT IS MEASURED BY MULTIPLE-CHOICE TESTS OF READING COMPREHENSION

Hannon, Sandra Furr
A VALIDATION OF THE MENZEL BRANCH COUNT AS A MEASURE OF SYNTACTIC COMPLEXITY

LoVano, Gloria Maria
THE RELATIONSHIP OF SELF CONCEPT AND INTERSPERSED QUESTIONS TO AID COMPREHENSION IN THE PROCESSING OF TEXTUAL MATERIALS

McGee, Lea M.
TRANSFER FEATURE THEORY APPLIED TO CLOZE PROCEDURES: EFFECTS ON COMPREHENSION AND RECALL OF GOOD AND POOR READERS

Moore, Sue Gafford
PROMOTION-RETENTION: A COMPARISON OF FIFTH GRADE STUDENTS' READING AND MATH SCORES ON THE FLORIDA STATEWIDE ASSESSMENT TEST

Nolting, L. Wray
RELATIONSHIPS AMONG A MEASURE OF AUDITORY AND VISUAL SEGMENTATION AND STANDARDIZED AND INFORMAL MEASURES OF READING READINESS

Owen, Jo Anne
THE DEVELOPMENT AND EVALUATION OF A WORDS-IN-CONTEXT TEST TO DETERMINE INSTRUCTIONAL READING LEVELS

Rauch, Margaret Marie
THE RELATIONSHIPS BETWEEN SELECTED DISCOURSE ANALYSIS TASKS PREVIEWING STRATEGIES AND SUBSEQUENT FIFTH AND SIXTH GRADE STUDENTS' PERFORMANCE ON MEASURES OF READING COMPREHENSION

Rowe, M. Elise Murphy
AN EXPERIMENT FOR DETERMINING THE MAIN IDEA IN SECONDARY SCHOOL READING: THE SUBSKILL METHOD VERSUS THE MODIFIED LANGUAGE EXPERIENCE METHOD

Shanahan, Timothy Edward
A CANONICAL CORRELATIONAL ANALYSIS OF THE READING-WRITING RELATIONSHIP: AN EXPLORATORY INVESTIGATION

Smith, Karen Gail
THE DEVELOPMENT AND ANALYSIS OF A CRITERION-REFERENCED TEST OF TEXTBOOK READING AND VOCABULARY DISCRIMINATION FOR UNIVERSITY FRESHMEN WITH REFERENCE TO THEIR ACT COMPOSITE SCORES

Smith, Virgil Dempsy
AN INVESTIGATION OF THE AURAL CLOZE PROCEDURE AS A MEASURE OF LISTENING COMPREHENSION

Summers, Pamela French
THE RELATIONSHIP BETWEEN DEMONSTRATED STORY GRAMMAR USAGE BY THIRD GRADERS AND THEIR SCORES ON SELECTED READING COMPREHENSION TESTS

Turner, Gladys H.
OPAL READING ERRORS OF FIFTH GRADE STUDENTS

Winters, Lynn
THE EFFECTS OF DIFFERING RESPONSE CRITERIA ON THE ASSESSMENT OF WRITING COMPETENCE

COMPARISON OF ORAL READING, SILENT READING, AND LISTENING COMPREHENSION OF THIRD GRADE STUDENTS ASSESSED BY GATES-MacGINITIE READING TEST

Order No. 8024235

AL-DAHIRY, SALEEM ABDUL ABBAS YASIN, PH.D. *Kansas State University*, 1980. 152pp.

The purpose of this study was to determine the effect upon comprehension by reading silently, by listening to reading, and by reading orally in a typical third grade, with control over intersubject variation and variation from differences in difficulty of reading material. It was hoped that the results from this study would apply to third grade instructional strategies.

Comprehension was measured by the Gates-MacGinitie Reading Comprehension Test, Level C, Form 2 (1978), which contains twenty-two different passages each with both a literal and an inferential question. These passages were randomly divided into groups of three with approximately equal difficulty. Each group was then divided, by means of random assignment, into groups of three selections--A, B, or C. Selections A and C had seven passages with 14 questions for each, and selection B had eight passages with 16 questions.

The experimental population for this investigation was taken from a Midwestern military post elementary school. The sample consisted of 60 third grade boys and girls, chosen by random selection and random assignment, and divided into three third grade groups of 20 children each.

Use of the Posttest-Only Control Group Modification Design, together with unbiased assignment of subjects to conditions provided suitable control for sources of internal validity.

The three groups read silently, listened to reading, and read orally in different order a series of three selections. Group one read selection A silently, listened to selection B, and read selection C orally. Group two read selection B silently, listened to selection C, and read selection A orally. Group three read selection C silently, listened to selection A, and read selection B orally. Time was kept constant for the three modes of learning.

The hypothesis to be tested was stated as a null hypothesis: There was no significant difference in comprehension effectiveness resulting from the three modes of reading the passages.

Data analysis was made through the One-Way Analysis of Variance and General Linear Models Procedure and Tukey's (HSD) Test. Significance level for the investigation was established at .05.

The results for the null hypothesis (all three groups in each mode considered together) tested for significance on the basis of *Total Mode Means* showed first, a significant difference between silent and oral comprehension in favor of oral comprehension; second, listening comprehension significantly superior over silent comprehension; third, no significant difference between listening and oral comprehension.

When analysis of variance was used to test for significance between groups in each mode on the basis of *comprehension scores*, the results showed no significant comprehension differences between the groups in silent reading, listening, or in oral reading.

Testing for significance *within groups* by General Linear Models Procedure and Tukey's (HSD) Test showed the following: Group I: no significant differences between silent mode and listening, silent and oral, or listening and oral; but listening was higher than both silent and oral comprehension; Group II: no significant differences within the subjects in three modes, with oral slightly higher than listening and silent comprehension; Group III: the oral mean significantly higher than listening and silent means which were not significantly different from each other.

Finally, no significant interaction was found between groups by treatments.

In conclusion, the sources of variation were controlled by random selection and assignment of third grade subjects and by giving the three selections to all three groups in rotating order to eliminate differences from difficulty in reading material. This showed that the independent variables are the three modes of learning--silent, listening, and oral reading--with oral reading leading to the highest comprehension and silent reading the lowest in these third grade groups which read the same material in counterbalance order.

THE EFFECTS OF TEST-TAKING INSTRUCTION VS. PRACTICE WITHOUT INSTRUCTION ON THE EXAMINATION RESPONSES OF COLLEGE FRESHMEN TO MULTIPLE-CHOICE, OPEN-ENDED, AND CLOZE TYPE QUESTIONS

Order No. 8016288

BERGMAN, IRWIN B., ED.D. *Hofstra University*, 1980. 136pp.

This study was conducted: (1) to determine the effectiveness of test-taking instruction in which nonproficient college readers were provided instruction in how to respond to questions of the following types: multiple-choice, open-ended, and cloze; and (2) to determine if a significant difference exists between test results achieved by the groups given test-taking instruction (Instructional Group) when compared with groups given no test-taking instruction and not afforded the opportunity to practice on the instrument (Control Group), and with groups which are permitted to practice on various test selections but are given no test-taking instruction (Practice Group).

The subjects of the study were 349 students enrolled in eighteen reading and study skills improvement classes at Queensborough Community College in Bayside, New York. The subjects were randomly divided into Instructional, Practice, and Control groups. Each group was pre-tested with a multiple-choice, open-ended, or cloze format of the Nelson-Denny Reading Test, Form B. Over a period of approximately two weeks, the Instructional group received formal test-taking instruction, the Practice group was permitted to practice test-taking without formal instruction, and the Control group was given work which was not related to test-taking. All groups were then post-tested with Form A of the Nelson-Denny Reading Test.

Post-test scores indicated the following: (1) the cloze Instructional group scored significantly higher at more than the .05 level that the cloze Practice and cloze Control groups; (2) no statistically significant differences were found to exist between the multiple-choice post-test scores or between the open-ended post-test scores of any of the groups; (3) the multiple-choice Instructional group attempted significantly more answers, at more than the .01 level, on the post-test instrument than did the corresponding Control group; (4) the cloze Instructional group attempted significantly more answers on the post-test than either the Practice or the Control groups; (5) no significant differences were found in any category between the Practice and Control groups.

The major conclusions, which were made on the basis of the findings of the investigation and within the limitations of this study, are as follows: (1) Test-taking instruction for standardized multiple-choice tests and for open-ended tests did not appear to be an effective technique for improving performance on a standardized reading comprehension test. (2) Test-taking instruction for cloze tests appeared to be a fruitful technique for improving the performance of students in the completing of cloze responses on passages of a standardized reading comprehension test. (3) Test-taking practice without formal instruction did not appear to be a fruitful technique for improving the performance of students on a multiple-choice, open-ended, or cloze format of standardized reading comprehension tests. Any gains made as a result of practice effect were not statistically significant. It was concluded that practice on a particular test format unaccompanied by formal instruction was not sufficient in itself to effectuate significant student gains in test performance.

Most of the studies in the related research found test-taking instruction more efficacious than the results of this study suggest. The results of this study may be generalized only to the non-proficient college reader who is being assessed in standardized tests of skills such as reading rather than content.

ELECTRODERMAL RESPONSES AND INFORMAL READING CRITERIA

Order No. 8016510

BONDS, CHARLES WESLEY, ED.D. *The University of Florida*, 1979. 101pp. Chairman: William R. Powell

The informal reading inventory appears to be the best instrument available for determining the independent, instructional, and frustration reading levels. However, the criteria now in vogue for determining the levels must be scrutinized more critically. The rationale for the criteria used to determine levels as advocated by Betts is grounded on assumptions and wide spread usage rather than empirical research. Much of the controversy concerning the inventory is not whether it is a useful tool but the validity of the criteria used to determine the various levels. Physiological measures of reading performance could be used to circumvent the dilemma of appropriate criteria.

The primary purpose of this investigation was to determine if there was a difference between the attained mean galvanic skin responses of elementary school subjects according to readability clusters. Clusters refer to the attained readability level of the passage read by subjects. Clusters were designated as follow: Cluster I, levels two and below; Cluster II, levels 3-5; and Cluster III, 6-8.

The subjects for this investigation were 40 boys and girls distributed as follow: second graders, 5 males and 7 females; fourth graders, 9 males and 12 females: and sixth graders, 6 males and 7 females.

Data used in the study were collected while the subjects' ring and index fingers of the right hand were attached to electrodes of a Lafayette Instrument Company galvanometer. As the subjects read passages of increasing difficulty from an informal reading inventory, galvanic skin responses were recorded. The galvanic skin responses were recorded at the beginning and end of each line of the selections read by the subjects. During the answering of comprehension questions, the magnitude of the deflection of the galvanometer needle was recorded and used to calculate mean galvanic skin responses measures while the subjects answered comprehension questions relating to the passage read orally.

The results of this study indicate that as the difficulty level of the materials increases, the subjects make fewer and fewer word recognition errors. At the lower levels of reading, the reader is able to endure more word recognition errors and maintain adequate comprehension. The Betts criteria of uniform application of error percentages was not evident in this investigation. No two reading levels or clusters of performances were the same. Each cluster produced its own unique word recognition percentages. The percentages achieved by this group of subjects were not equivalent to Powell's criteria. However, they showed a definite decrease in the number of word recognition errors as the readability level of the materials increased from the lower to the higher levels. Betts' uniform criteria were not evident.

This investigation also revealed that there are no significant differences between the mean galvanic skin responses of subjects at both the instructional and frustration levels while reading orally. Significant differences were found when the subjects responded to questions at the instructional level. At levels 3-5, readers became more excited when answering comprehension questions than subjects attaining a readability level of 2 and below and at the levels 6-8.

EFFECTS OF A RATE EMPHASIS ON THE READING PERFORMANCE OF REFLECTIVE AND IMPULSIVE SIXTH-GRADE MALES AND FEMALES

Order No. 8015787

BRINZER, RAYMOND JOSEPH, PH.D. *The Pennsylvania State University*, 1980. 137pp. Advisers: Emery P. Bliesmer and Frances Fairchild

The purpose of this study was to determine the short-term effects of a reading rate emphasis upon reflective and impulsive male and female student reading performance.

Matching Familiar Figures (MFF) tests were administered to 180 sixth-grade students. Participating students were then categorized reflective, impulsive, or unclassified, and assigned to study groups based on their SRA reading test scores, sex, and cognitive style. Teachers were randomly assigned to the treatment groups.

Gates-MacGinitie Reading Tests (GMRT) were administered. Students were assigned reading material levels and the study commenced. The manipulated treatment variable was reading rate. Reading rate was manipulated via teacher monitored drill-type practice for speed treatment (fast) and non-speed treatment (careful) reading groups over a 20-day experimental period. Data necessary to gauge treatment effects were obtained when GMRT posttest administration occurrred.

The experimental design was mixed. Tests of hypotheses at the .05 significance level were conducted via separate three-factor ANOVR (Analysis of Variance with Repeated Measures) for sex, and for style, while pairwise statistics were used for follow-up significant difference analyses.

The overall *F* tests revealed several significant differences. Pairwise comparisons of significant difference means revealed a significant within-subjects difference for female vocabulary growth ($p < .05$), and a significant between-subjects difference in comprehension for reflective students ($p < .05$).

Raw scores were further analyzed by a SAS (Statistical Analysis System) program in order to obtain additional descriptive data for post hoc analyses. Those analyses revealed comprehension declines for all groups analyzed except reflective students, who posted comprehension raw score increases.

STANDARDIZED READING TEST SCORES AND INFORMAL READING INVENTORY RESULTS WITH MISCUE ANALYSIS FOR ABOVE AND BELOW AVERAGE THIRD GRADE STUDENTS: A COMPARISON

Order No. 8028320

BUTTON, LINDA JEAN ORVIS, ED.D. *University of Northern Colorado*, 1980. 122pp.

Purpose of the Study. The purpose of this study was to compare grade equivalent reading scores obtained on the *Woodcock-Johnson Psycho-Educational Battery* Reading Cluster by (a) third grade students classified as below average in reading and (b) third grade students classified as above average in reading, with the instructional reading levels of the students ascertained using an informal reading inventory. In addition, reading miscues for each student were analyzed according to a modified miscue analysis procedure to determine whether or not above average readers and below average readers at the third grade level used similar reading strategies a majority of the time.

The final population sample consisted of forty students: twenty below average and twenty above average third grade students in reading.

Procedures. The *Woodcock-Johnson Psycho-Educational Battery* Reading Cluster was administered to the third grade students. The functional reading levels for each student were identified by administering an informal reading inventory and applying Betts' criteria. The students' oral reading was coded and individually analyzed according to a modified miscue analysis procedure to determine what reading strategies were used.

Analysis of the Data. The means, standard deviations, and ranges were computed for scores attained on each assessment device for each student group. A correlation coefficient was calculated for each group to show the relationship between the scores of the standardized test and the instructional level of the informal reading inventory. A t-test for correlated samples was used to determine whether or not the results were significant at the .05 level using a two tailed test. Grade placement comparisons for individual students were made through descriptive, tabular and graphic analyses.

After each student's miscues were analyzed, dominant reading strategies were determined. The reading strategies for each group were compared through tabular and descriptive analyses.

Conclusions. On the basis of the findings and the limitations imposed by the study, the following conclusions were drawn: (1) The *Woodcock-Johnson Psycho-Educational Battery* Reading Cluster grade score mean underestimated the mean instructional reading level of third grade students classified as below average in reading. (2) The *Woodcock-Johnson Psycho-Educational Battery* Reading Cluster mean did approximate the mean instructional reading level of third grade students classified as above average in reading. (3) Results from the standardized test and the informal reading inventory were moderately related for below average third grade students. However, the *Woodcock-Johnson* Reading Cluster underestimated the below average students' instructional reading level by one half year. (4) Results from the *Woodcock-Johnson* Reading Cluster and the instructional level of the informal reading inventory were minimally related for the above average third grade students due to extreme variations in individual scores. (5) Reading strategies for below average students in reading incorporated grapho-phonic and semantic syntactic cues in nearly equal proportions while above average students in reading in all cases relied on semantic-syntactic cues as a dominant reading strategy and appeared to rely very little on grapho-phonic cues. As below average students approached grade level in reading, they relied less on grapho-phonic cues and more on semantic-syntactic cues as a dominant reading strategy.

A STUDY OF THE RELATIONSHIP BETWEEN THE CLOZE TEST AND A HIERARCHICAL MODEL OF READING COMPREHENSION SKILLS

CALICA, DIANA KRAKOWER, PH.D. *The University of Chicago*, 1980

The cloze procedure is an objective, inexpensive and easily constructed instrument for measuring literacy. A critical problem associated with its use has been the inability to interpret scores on this test in terms of general reading processes. The purpose of this study was to investigate the relationship between scores on literal reading comprehension skills tests and scores on the cloze test.

Three literal reading comprehension skills (*i.e.* word comprehension, sentence comprehension, and intersentence comprehension) were defined on the basis of a linguistic theory. A rationale was developed for hypothesizing that the three skills form a hierarchical relationship, with proficiency in using a higher level skill depending upon the extent to which an individual has mastered a low level skill.

Hypotheses linking performance on tests of these reading skills with performance on the cloze test presented the following notion: what a cloze test might index for a particular individual was the level of proficiency (s)he had achieved in the skills hierarchy. It was specifically proposed that (1) a given cloze score would reflect a higher mean level of performance on a simple skill than on a more complex skill, and (2) as cloze scores rose, they would index a subject's ability to apply increasingly complex skills to a material.

The investigation involved administering a cloze test to each of the fourth grade subjects. Three days later the subject received a passage and accompanying skills test that corresponded to the cloze test (s)he had already taken.

Of major concern in the study were: (1) the development of testing instruments that would measure in an objective fashion a subject's ability to apply a specific comprehension skill, and (2) a non-subjective means of scoring the open-ended answers to these tests. The considerations that went into constructing algorithms for the generation of the skills test items and a grading system are presented in the report.

The experimental results were as follows: (1) the mean scores of the skills tests decreased as the skill being tested increased in complexity; (2) within the range of 95% of the scores, any given cloze score reflected a higher mean level of performance on a simple skill than on a more complex one; (3) as cloze scores rose from 0% to 100%, they relfected mastery of the vocabulary in a material first, mastery of sentence comprehension next, and mastery of intersentence comprehension last.

A COMPARISON OF A STANDARDIZED GROUP READING TEST AND AN INFORMAL READING INVENTORY IN ESTIMATING AN INSTRUCTIONAL READING LEVEL

Order No. 8028323

CLEMENT, BETTY ANN WAIDLICH, ED.D. *University of Northern Colorado*, 1980. 137pp.

Statement of the Problem. The purpose of this study was to compare the instructional reading levels estimated from the administration of the *Metropolitan Achievement Test*, and an informal reading inventory for fifth grade students classified as below average, average, and above average readers.

Procedures. Three groups, of thirty fifth grade students each, were included in the study. The Reading Comprehension portion of the *Metropolitan Achievement Test* was administered in a group setting. The students were classified as below average, average or above average based on grade equivalent scores from the *Metropolitan Achievement Test*. The instructional reading level was also obtained from the *Metropolitan Achievement Test*.

The *Ekwall Reading Inventory* was administered orally to each student and an instructional reading level was estimated using Betts' criteria. The resulting instructional reading levels from the *Metropolitan Achievement Test* and the *Ekwall Reading Inventory* were compared.

Analysis of the Data. Tables were prepared to report the descriptive information obtained from the schools and the *Metropolitan Achievement Test* and informal inventory instructional reading levels. Means and standard deviations for the *Metropolitan Achievement Test* and informal reading inventory instructional reading levels and the Pearson product-moment correlation coefficients between the scores of the two tests were calculated for the below average, average, above average readers and for the combined group. A t-test for related samples was used to determine the significance of the results at the .05 level.

Conclusions.

Research Question. The *Metropolitan Achievement Test* and the informal reading inventory produced quite different instructional reading levels for *individuals* who are below average, average, or above average readers. Using the *Metropolitan Achievement Test* instructional reading levels alone to place fifth grade students in materials or groups for instruction may result in inaccurate placement for a majority of the students.

Hypotheses I and II. For the below average *group* the two instruments on the average produced substantially different reading levels, with the *Metropolitan Achievement Test* placing students approximately one grade level higher than the informal reading inventory.

Even though the average group tended to score higher on the *Metropolitan Achievement Test* than on the informal reading inventory and the above average group tended to score lower on the *Metropolitan Achievement Test* than the informal reading inventory, the average differences for these groups, and for the total group, were not substantial. However, the variation in individual scores above and below the informal reading inventory instructional reading level was sufficient to cause most students to be placed at a frustration reading level or a level too easy for them.

Since the primary concern in the classroom and in this study was for accurate placement of individuals and not determining the average reading levels of groups, as considered in the two hypotheses, the *Metropolitan Achievement Test* was judged as ineffective in accurately placing individuals at appropriate instructional reading levels.

AN ANALYSIS OF FOURTH AND SIXTH GRADE READER PERFORMANCE USING CLOZE TESTS, GROUP READING INVENTORIES, AN INFORMAL READING INVENTORY, AND A STANDARDIZED READING TEST

Order No. 8020381

DAVIDSON, EMMA SUE, ED.D. *Memphis State University*, 1980. 92pp. Major Professor: Dr. Thomas A. Rakes

Overview. This study compared results from an informal reading inventory (IRI), a standardized reading test (SRA), a group reading inventory (GRI), and a cloze test when used to measure reading performance in grades 4 and 6. Comparisons were made using test scores and reading levels. Pupil performance was compared in social studies and in science areas for the group reading inventories and the cloze tests for grades 4 and 6.

Procedure. A population of 624 students in grades 4 and 6 from 13 schools in northeast Arkansas and southeast Missouri was administered the SRA achievement test, a science and social studies group reading inventory, and 50 students were administered an informal reading inventory (April and May, 1979). Correlations between the variables were computed and differences between correlations were tested by Fisher's *z* test. To further explore the relations among the variables, factor analysis was used. When comparing reading levels on the four reading tests, Kendall's Tau B was used to analyze the data. A significance test on the differences between correlation coefficients was performed. An analysis of variance was performed comparing grades 4 and 6 on the cloze science and cloze social studies and the GRI science and GRI social studies. Specifically, a 2 (grade) x 2 (content: science vs. social studies) ANOVA with repeated measures was employed.

Findings. When analyzing test scores for grades 4 and 6, the strongest relationships were between the cloze and GRI test results. The weakest were correlations of the IRI to the cloze and GRI tests. SRA and IRI correlated significantly. In the factor analysis, SRA and IRI loaded on Factor 2, which was labeled "general reading." Cloze science, cloze social studies, GRI science and GRI social studies loaded heavily on Factor 1, which was labeled "content reading and concept processing." Kendall's Tau B was computed to determine the similarity of tests when analyzing reading level placement. For both grades 4 and 6, the inter-correlations were generally low, with the highest relation between the SRA and IRI. The cloze and GRI tests correlated moderately and were found to be more highly correlated with one another than with SRA. An unexpected finding was the negative correlations between the IRI and cloze tests and IRI and GRI tests. An

ANOVA was used to determine the effects of grade and subjects. This analysis of variance showed that there was a main effect of grade and that there was interaction between grade and subject.

Conclusions. The following conclusions were based on the findings of this study: (1) the diagnosis of reading ability needs to go beyond the measurement of general reading if the user intends to examine the reader's ability to apply the process of reading content material in grades 4 and 6; (2) students may be performing poorly in content subjects not because they lack general reading ability, but because they lack the ability to apply the necessary reading skills for content reading; (3) since the cloze procedure has been validated as a useful measurement procedure in reading and the data from this study indicates a strong observed relationship between GRI and cloze test scores, there is an indication of a degree of support for using the GRI in grades 4 and 6; and (4) results did not support the contention that standardized test results produce an inflated reading score when compared to reader achievement on an IRI.

Recommendations. There are several follow-up studies suggested by findings in this study: (1) The investigation should be replicated using the same tests with the same and different grade levels in an effort to verify or refute the findings in this study. (2) The investigation should be conducted using an IRI constructed from subject area materials. (3) Further research is essential to confirm that the findings in this research are of general significance.

RELATIONSHIP OF READING ACHIEVEMENT TO CERTAIN SELECTED TESTS AND NEWLY CONSTRUCTED MEASURES OF PSYCHOLOGICAL PROCESSES

Order No. 8020120

DE SOTO, JANET LOUISE, ED.D. *The Johns Hopkins University*, 1980. 175pp.

This study is an investigation of reading achievement and its relationship to a broadly representative collection of commonly used tests as well as newly constructed measures of psychological processes. Existing instruments were used to evaluate the following processes: short-term memory (evaluated with Visual Attention Span for Objects, Auditory Attention Span for Unrelated Words, Auditory Attention Span for Related Syllables), associative learning (evaluated with the Van Wagenen Word Learning Test), and semantic association (evaluated with the Verbal Opposites test). Tasks were constructed to evaluate the remaining processes inasmuch as standardized tests did not seem to be available for their evaluation. These processes and tasks included: encoding speed for pictures and words (evaluated with Naming Pictures and Reading Words), recoding unknown words (evaluated with the Recoding Task and Correct Consonants Recoded), and automaticity of processing known words (evaluated with Automatic Word Processing).

Two groups of fourth-grade students were selected: (a) 67 achieving readers who scored at the 66th percentile or above on the reading comprehension subtest of the California Achievement Test and (b) 67 nonachieving readers who scored at the 34th percentile or below on the same test. The ten instruments were administered individually to the students in both samples.

All of the instruments, with the exception of Automatic Word Processing, were significantly related to reading achievement, and their use in evaluating deficiencies in psychological processing abilities associated with reading achievement was supported. When nonverbal IQ (as measured by the Cognitive Abilities Test) was controlled for by means of the analysis of covariance, achieving readers continued to demonstrate performance superior to the non-achieving readers on all the measures except Automatic Word Processing and Auditory Attention Span for Unrelated Words. This finding further supports their use in evaluating psychological processing abilities associated with reading achievement, except in the case of the Auditory Attention Span for Unrelated Words, for which caution in indicated. A stepwise multiple regression analysis indicated that the combination of psychological measures that most efficiently explained a large proportion of the variance of reading achievement was the Verbal Opposites test, Correct Consonants Recoded, and the Van Wagenen Word Learning Test.

The two tests that stood out with the highest correlations with reading comprehension were the Verbal Opposites test ($r = .71$) and the newly constructed Correct Consonants Recoded measure ($r = .60$). The Verbal Opposites test showed rather high correlations ($r = .35$ to $r = .55$) with the other psychological measures (except for Automatic Word Processing). These correlations suggested that the verbal processing abilities underlying performance on the Verbal Opposites test also influenced performance on the other psychological measures as well as reading achievement.

THE UTILITY OF SELECTED INTERACTION ANALYSIS SYSTEMS AS FEEDBACK TOOLS DURING THE READING INSTRUCTION PERIOD

Order No. 8015601

EDWARDS, KATHRYN INEZ, PH.D *Claremont Graduate School*, 1980. 219pp.

This study investigated the utility of five interaction analysis systems for studying verbal behaviors that communicate positive and/or negative expectations that teachers hold for students. Each system was selected on the basis of its potential for observing: (1) praise and (2) quality of questioning. Within the questioning category two subcategories were investigated -- the use of direct or closed teacher questioning and the use of open/indirect questioning. Data were reported on the findings of the systems in terms of inter and intragroup comparisons. However, the major thrust of the study was the investigation of the capability of each system to enable a researcher to observe specific components of classroom interaction. It was the methodology of research rather than the findings that were of primary importance here.

A review of former research has determined that teacher interactions with students perceived as "low achievers" are less supportive and less motivating than those interactions normally practiced with students perceived as high achievers. Teachers tend to ask high ability students questions that require the use of higher cognitive processes. Lower ability students are frequently asked questions that require answers that call on rote memory or immediate recall. The use of praise was also investigated. Teachers have been found to consistently praise the positive behaviors of high ability students, while frequently ignoring positive behaviors of low ability students. Low achievers are more likely to be the subjects of criticism than are high achieving students.

Two schools in the inner city Los Angeles area participated in the study. One fourth grade class in each school was tape-recorded on two occasions. The tapes for the high and low ability reading groups were selected for study. Five-minute segments were randomly selected from each of the tapes by an independent researcher for transcription. Each segment was analyzed according to the interaction analysis systems of Douglas Barnes, Arno Bellack, Ned Flanders, Jere Brophy and Thomas Good, and Bruce Joyce. All occurrences of praise interactions were coded as well as occurrences in the two subcategories of questioning. All information was analyzed within the structure of the particular system under consideration. Total scores for each system in each of the three categories were recorded on a summary page. These data were then subjected to a T-test to measure the degree of statistical significance.

The comparisons showed that for all groups studied, low achievement students received more praise than did high achievement students. The systems studied also revealed that the students in the top group were asked more direct questions than were students in the low group. The differences between the high group and the low group in the number of open questions asked was minute and no significant differences were discovered.

These findings, which were contrary to those of similar research studies were explained in terms of two factors: (1) comparability of the achievement range of students in this study to students in former studies and (2) the size and composition of the sample populations. Recommendations were made on the basis of the findings of the study. The recommendations have implications for the use of interaction analysis systems as feedback tools for teacher training programs. The systems provide a mirror for the teacher to obtain feedback about his own teaching behavior along the demensions of the particular system used. The systems are also useful as observational tools for supervisors. Finally, classroom observation instruments enable the teachers to perform "micro-research" on their own behavior in the classroom and thereby should lead to an increase in flexibility and general improvement in inter-relatonships with students. In this way the teacher can maximize his/her role in guiding pupils' learning.

AN ETHNOGRAPHIC INSTRUMENT TO ASSESS THE LINGUISTIC INTERACTIVE BEHAVIOR OF CLASSROOM GROUPS

Order No. 8013322

FOREST, ROBERT GERARD, ED.D. *Boston University School of Education*, 1980. 308pp Major Professor: Roselmina Indrisano

The purpose of this study was to devise an ethnographic instrument to assess linguistic interactive behavior in the classroom.

The study was developed in five stages. (1) The literature in group process was reviewed to serve two purposes: (a) Identification of the basic group process principles to be translated into linguistic behaviors; (b) Determination of the knowledge that the teacher/assessor must possess to interpret linguistic behavior. (2) A prototype linguistic instrument was designed from 31 group process principles, and was tested at kindergarten and grades 1, 3, 5, 7, and 11 to verify applicability. (3) Thorndike and Hagen's steps for systematizing the direct observation of behavior were followed to produce the refined instrument. (4) Ethnographic factors were applied to the development of the approaches recommended to administer and interpret the Linguistic Assessment Profile. (5) Interrater reliabilty analysis was employed to compare the results obtained by four assessors analyzing the same interactive session. The quantifiable data were analyzed through the computation of the Hoyt correlation coefficient computed from the analysis of variance data.

The following conclusions were reached:

Question 1. What are the basic principles underlying the group process? Outcome. A search of the literature revealed 31 group process principles, each cited by a minimum of three authoritative sources.

Question 2. In addition to the principles, what knowledges about group process theory must a teacher possess to facilitate understanding of the linguistic behavior that occurs in the classroom? Outcome. The classroom teacher must possess a repertoire of interactive skills that represent knowledges in the following areas: (1) the broad concerns of group functioning; (2) the stages of group development; (3) the specific concerns of group functioning

Question 3. What pertinent information regarding the description of the group process principles and the employment of ethnographic methods in the collection and analysis of data forms the linguistic and ethnographic basis for the instrument design? Outcome. Group process principles were translated into linguistic behaviors to become the basic and linguistic descriptors were compiled to define the parameters of the categories; and ethnographic factors influencing the participant/observer approach to linguistic analysis were identified to govern the procedures.

Question 4. How are the group process and ethnographic factors selected from the research and test analyses synthesized into a dynamic model for linguistic interactive assessment in the classroom? Outcome. Based on the research, two documents were designed for the collection and analysis of linguistic data: (1) The Assessment Profile Record was developed to record and quantify speech events. (2) The Linguistic Assessment Profile was developed to assign speech events to precise linguistic categories and network channels.

Question 5. How reliable is the ethnographic approach to linguistic assessment? Outcome. The resultant linguistic instrument is considered reliable in that its consistency between assessors was affirmed.

Summary. The conclusions support the construction of an ethnographic instrument to assess linguistic interactive performance in the classroom, beginning at grade 3. The profile, relying on research data gleaned from the literature in group process, ethnography of speaking, and theory and research based views of interaction, is dependent upon group problem-solving approaches and the teacher assuming a participant/observer role.

THE EFFECTS OF PASSAGE LENGTH AND READING MODE ON RETELLINGS AND QUESTION RESPONSES OF SIXTH GRADE CHILDREN

Order No. 8017800

GRANT, NORMA LOUISE, ED.D. *The University of Arizona*, 1980. 136pp. Director: Wilbur S. Ames

The purpose of this study was to investigate differences in oral and silent reading comprehension performance with passages of varying length through the use of both a questioning and a retelling format.

In analyzing reading comprehension performance scores, answers were sought for the following questions: (1) What differences between oral and silent reading comprehension performance can be identified with passages of varying length when question responses are the measure of comprehension? (2) What differences between oral and silent reading comprehension performance can be identified with passages of varying length when retelling scores are the measure of comprehension? (3) In either oral or silent reading, to what extent is the information sought by means of traditional questions supplied by the reader during a retelling activity?

To answer these questions, 80 sixth grade students from one public school located in a southwestern metropolitan area were randomly assigned to eight treatment groups. Subjects were of average ability as measured by a standardized test of silent reading achievement.

A story with a sixth grade readability level was considered the long passage. The first portion of this same story was considered the short passage.

All subjects completed an experimenter constructed instrument designed to reveal preferences for oral or silent reading in various situations.

The 40 subjects randomly assigned to the question treatment groups read the assigned passage, either long or short, in the assigned mode of presentation, either oral or silent, and responded to questions asked orally by the experimenter. The 40 subjects randomly assigned to the retelling treatment groups also read the assigned passage in the assigned mode of presentation and retold the story to the experimenter. Open-ended questions were asked by the experimenter in order to elicit additional information. Following the retelling, the questions used with the question treatment groups were asked of the subjects in the retelling treatment groups. Access to the story was not allowed during retelling or questioning. Finally, subjects in the retelling groups were asked to react to various vocabulary items when again shown the story.

Findings based on the statistical analysis of the data gathered in this study were as follows: (1) No significant differences were found between oral and silent reading comprehension of long and short passages when either questioning or retelling was the measure of reading comprehension performance. (2) Significant differences were found in performance between the questioning and retelling groups on questioning tasks when scores for the retelling groups were based only on information supplied during spontaneous retelling and open-ended probing. (3) Significant differences were found favoring a preference for silent over oral reading in various reading situations. (4) In the questioning condition, silent reading of passages was significantly faster than oral reading of passages. (5) No signincant differences were found in the words per second reading rate for passages in the retelling condition.

The findings of this study supported these conclusions: (1) Regardless of mode of comprehension assessment, students are able to perform equally well after reading orally or silently. (2) Varying the length of passages does not affect performance in reading comprehension tasks. (3) Different modes of comprehension assessment do not necessarily yield the same kinds of information about reading ability. (4) Students prefer silent reading as an activity but that preference is not reflected in superior silent reading performance. (5) Students, when aware of the mode of comprehension assessment to be used, may make differential adjustments in their rate of reading.

DETERMINING WHAT IS MEASURED BY MULTIPLE-CHOICE TESTS OF READING COMPREHENSION

HAERTEL, EDWARD HENRY, PH.D. *The University of Chicago*, 1980.

This study explored the use of qualitative models with misclassifications to describe reading comprehension test response data. Based on a review of reading comprehension literature, nine hypothesized skills were described, which might be required by some but not all multiple-choice reading comprehension test items. These included comprehension of embedded propositions, anaphora, and propositions conveyed by inter-sentence syntax; locating the tested proposition in the passage provided when the linguistic focus of the question stem was altered, when the question reflected semantic substitutions for terms in the passage, and when certain types of more complex transformations had occurred; solving items requiring inferences based strictly on passage content or inferences involving additional information; and determining that best name for or main idea of the passage. Single items often required several skills. A coding scheme was developed to determine which skills reading comprehension items required, based on linguistic features of the items, and the inter-rater reliability of this coding guide was established.

Response data from large, nationally representative samples of upper elementary children were obtained from two recent federally sponsored studies, and hypothesized skill requirements were determined for all items on standardized reading comprehension tests used in these studies. Sets of five items requiring different subsets of selected combinations of the nine skills were chosen for analysis. For each item set, probabilities of each possible pattern of correct and incorrect responses were then computed.

In modeling these response pattern probabilities, each student was assumed to conform to exactly one of two or more latent states, determined solely on the basis of the skills he or she possessed. If the subsets of skills items required and students possessed were the sole determinant of correct and incorrect responses, then for students in any given latent state, only one pattern of correct and incorrect responses would occur. Since in fact students lacking requisite skills might nonetheless respond correctly to an item and conversely, misclassification parameters were included, giving the probabilities of correct and of incorrect responses to each item conditional upon conformity to each possible latent state. Parameterization of misclassifications was studied in a series of preliminary analyses.

Once a complete model was described, including latent states and misclassification probabilities, parameters representing the proportions in each latent state and parameters representing the various misclassification probabilities were estimated simultaneously by the method of maximum likelihood. Standardized residuals and likelihood ratio chi squares were examined to determine goodness of fit. Difference chi squares were used to test hypotheses concerning the existence of the distinct skills and hierarchical relationships among skills.

None of the hypothesized skill distinctions was confirmed. However, a simple model involving only two latent states was found to fit response data extremely well, as indicated by fit statistics and raw and standardized residuals. According to this two-state model, roughly half of all fourth, fifth, or sixth grade children lack the skills required to solve any items, while the remainder possess skills required to solve all of the items. Each item is characterized by two distinct misclassification parameters, giving the probabilities of correct responses by students in the "cannot solve" and "can solve" latent states, respectively.

By methods used in this study, it is possible to distinguish the skills necessary to solve reading comprehension items from ancillary requirements for information (including vocabulary) prerequisite to the application of those skills. Findings indicate that much of the reliable variance in reading comprehension test performance may be a function not of students' skills, but of the ancillary requirements of individual items.

A VALIDATION OF THE MENZEL BRANCH COUNT AS A MEASURE OF SYNTACTIC COMPLEXITY Order No. 8014126

HANNON, SANDRA FURR, PH.D. *The Florida State University*, 1979. 279pp. Major Professors: Kellogg W. Hunt and Peter Menzel

A validation study was undertaken to determine whether or not the Branch Count, as restricted by Peter Menzel, is a valid measure of syntactic length and complexity in written language. In answer to the demand for quantitative measures of oral and written language acquisition, Menzel worked out a set of guidelines for drawing tree diagrams over the surface structure of syntactic units and then counting branches to determine both length and complexity of the units.

The Branch Count was validated by correlating its results with those of two established means of studying written language--Kellogg Hunt's words per T-unit count and John Bormuth's cloze measures--when these instruments were used to analyze thirty-two written passages. Moreover, the Branch Count results were also compared with a lesser known scale used by Walter Loban to study elaboration in communication units.

The analyses and the statistical treatment carried out to compare the Branch Count with the other measures of syntactic length and complexity indicate the following: (1) There is a high degree of correlation (.991) between the mean branch count per T-unit and the mean word count per T-unit in the passages used. Regression equations based on the results of analyzing these samples make it possible to convert word per T-unit counts to branch counts per T-unit and vice versa. (2) A high degree of correlation (.970) also exists between the mean branch count per T-unit and the mean number of elaboration index points per T-unit for the passages. Again regression equations allow for conversion of one count to the other. (3) The branch count per T-unit, analyzed as a function of both word count per T-unit and elaboration index points per T-unit, shows a positive correlation of .992 with these other measures in multiple correlation. When the mean branch count per T-unit is treated as a function of the sum of the other two counts, the coefficient of simple correlation is .989. (4) The Branch Count comparison with Bormuth's various cloze measures shows that: (a) The branch count per T-unit has a negative correlation of .852 with cloze means established by administering cloze tests on the thirty-two passages. A negative correlation is to be expected since the more syntactically complex the structure, the lower the expected cloze mean whereas the branch count is expected to rise with increased complexity. Regression equations again allow for converting one measure into an estimate of the other. (b) The correlation coefficient for the relationship between the branch counts per T-unit and the formula estimated cloze means is –.831. Regressions are given for this relationship, too. (c) The branch count per sentence treated as a function of both letters per sentence and letters per word has a high multiple correlation with these measures. The coefficient of multiple correlation is a positive .988. Regression equations make it possible to convert a branch count per sentence into an estimate of letters per sentence and letters per word so that a researcher can use these estimates with Bormuth's table based on cloze means to estimate reading levels for passages.

These results lead to the conclusions that Menzel's Branch count is a valid instrument for measuring syntactic development and that it does measure both length and complexity of syntactic units. However, the Branch Count is judged to more effective for measuring syntactic maturity in composing written language rather than in reading written language.

THE RELATIONSHIP OF SELF CONCEPT AND INTERSPERSED QUESTIONS TO AID COMPREHENSION IN THE PROCESSING OF TEXTUAL MATERIALS Order No. 8021702

LO VANO, GLORIA MARIA, PH.D. *Case Western Reserve University*, 1980. 142pp.

This study was designed to examine a sample of sixth-grade students for the relationships that exist between various subareas of self concept, meaningful learning and rote learning post questions, along with incidental learning.

To provide an understanding of the relationship between self concept and comprehension, the dependent variables for this study were: intentional statements, intentional concepts, intentional subsumed statements, incidental statements and incidental concepts.

The independent variables were self concept composite, self acceptance, social confidence and school affiliation. The independent variables of the superior and average comprehender were used to assess the interspersed meaningful learning and rote learning post questions and the incidental learning.

The subjects were 217 sixth grade parochial elementary students in the Cleveland, Ohio area. The subjects were tested on a 820 word reading passage. Each of the subjects in the experimental condition received either a meaningful learning post question or a rote learning post question throughout the passage. Each of the subjects in the control condition received the same stimulus materials with no questions throughout the passage. Materials were constructed at the sixth-grade reading level and were presented in random order. All of the students filled in a self report questionnaire which assessed the various dimensions of self concept.

The influence and/or relationship of self concept to reading performance was indicated by the results obtained with the self concept subscales interacting with the reading comprehension levels and the various reading measures. Close examination of the superior and average comprehenders' means showed a consistent difference between the two reading levels. It appeared throughout the entire experiment that the superior comprehender attained the higher means. The results obtained in this study support, in part, Ausubel's work in that the superior comprehender when presented with an advance organizer yielded greater recall than the rest of the sample population. Also, the rote learning condition significantly helped the average comprehender attain the higher means.

In general, the findings suggest that the relationship between self concept and reading comprehension exists. The average comprehender with the high self concept attained higher reading means than the average comprehender with the low self concept. These findings represent a possible explanation for test performance differences and if so, suggest the interaction of affective and cognitive determiners. It was, therefore, recommended that the relationship between comprehensive reading to learn and self concept should be studied within the framework of test performance such as free recall and multiple choice.

TRANSFER FEATURE THEORY APPLIED TO CLOZE PROCEDURES: EFFECTS ON COMPREHENSION AND RECALL OF GOOD AND POOR READERS Order No. 8028153

McGee, Lea M., Ed.D. *Virginia Polytechnic Institute and State University*, 1980. 173pp.

Based on transfer feature theory and spread of encoding explanatory frameworks, this study investigated comprehension strategies children use when processing written discourse on a meaningful level. Specifically, effects on comprehension and recall of a processing strategy requiring children to generate responses to cloze deletions was examined. In addition, this study investigated good and poor readers' ability to recall main ideas and details and the effects of anticipating cloze deletions which forces equal attention to main ideas and details on their recall of those ideas. Finally, the subsequent effects of this processing strategy on good and poor readers' immediate and delayed recall was investigated.

Forty good third grade and forty poor fifth grade readers recalled an easy cloze, difficult cloze, and normal passage. Twenty good and twenty poor readers retold each passage immediately after reading and after a one-week delay. The remaining readers retold each passage only after a one-week delay. Retellings were scored for proportion of total, main, and detail idea units substantively recalled according to analysis of discourse procedures.

Three factorial designs were used: three-way analysis of variance (passage x reader x recall) on total idea immediate and delayed recall, two-way analysis (passage x reader) on total idea delayed-only recall, and four-way analysis (passage x reader x recall x idea unit type) on main idea and detail idea unit immediate and delayed recall. Passage, recall, and idea unit type were repeated factors. Duncan's tests were used to examine post hoc mean differences.

Analyses yielded no main effect for passages in immediate and delayed recall. A passage x recall interaction ($p < .05$) revealed that easy cloze and normal passages were better recalled than difficult cloze passages in immediate recall. No mean differences across passages were found in delayed and delayed-only recall. Predicted passage recall facilitation due to increased spread of encoding was not supported.

Poor fifth grade readers recalled more idea units overall than good third grade readers ($p < .05$). In normal reading poor readers recalled more main ideas than good readers ($p < .05$), but good and poor readers recalled details equally well. Cloze passages did not influence good and poor readers' recall of main ideas or details. It was concluded that both good and poor readers attend to main ideas and details. Although good and poor readers recalled significantly less in delayed recall than in immediate recall ($p < .05$), in delayed recall poor readers recalled more than good readers. However, in delayed-only recall poor readers recalled only as well as good readers. These results were interpreted in light of possible good and poor reader processing differences. Suggestions for further research investigating these differences were proposed.

Additional findings indicated that good readers recalled proportionally more detail ideas than main ideas while poor readers recalled main ideas and details equally well. A possible developmental trend in children's ability to recall main ideas over details was suggested. This trend was interpreted within a structural processing explanatory framework.

PROMOTION-RETENTION: A COMPARISON OF FIFTH GRADE STUDENTS' READING AND MATH SCORES ON THE FLORIDA STATEWIDE ASSESSMENT TEST Order No. 8014140

Moore, Sue Gafford, Ph.D. *The Florida State University*, 1979. 127pp. Major Professor: C. Glennon Rowell

The objectives stated for this investigation were: (1) to determine if retaining a student one or more years in the K-5 grades makes any difference in his/her achievement in reading on the statewide assessment test; (2) to determine if retaining a student one or more years in the K-5 grades makes any difference in his/her achievement in math on the statewide assessment test; (3) to determine if school entry age of retained students makes any difference in fifth grade students' performance on the statewide assessment test; (4) to compare black and white retained students' fifth grade math scores on the statewide assessment test to determine if there is any difference; (5) to compare the math scores on the statewide assessment test of fifth grade black students with fifth grade white students who have not been retained; (6) to compare black and white retained students' reading test scores to determine if there is any difference; (7) to compare the reading test scores of black and white students who have not been retained to determine if there is any difference; and (8) to compare test scores in the areas of math and reading between fifth grade boys and girls who were retained and those who were promoted.

The population was composed of fifth grade students in eighteen elementary schools in Bay County, Florida. A stratified random sample was used to select students from the total population of fifth grade students. An effort was made to balance the population as to race, sex, retained and not retained students.

Data were collected from reading and math sections of the fifth grade Florida Statewide Assessment Test and students' permanent records. The t Test of significance of difference between means was used to analyze the data.

The conclusions from the study were: (1) Students who were retained scored lower in reading on the Florida Statewide Assessment Test than those who were not retained. (2) Students who were retained scored lower in math on the Florida Statewide Assessment Test than those who were not retained. (3) School entry age did not appear to make any difference in students' performance on the Statewide Assessment Test. (4) White retained students scored higher than black retained students on the math and reading sections of the Statewide Assessment Test. (5) White students who had not been retained scored higher on the math and reading sections of the Statewide Assessment Test than black students who had not been retained. (6) Girls who had not been retained scored higher than boys who had not been retained in the area of reading. (7) Retained girls scored higher than retained boys in the area of reading. (8) There was no significant difference in the math scores of retained boys and girls on the Florida Statewide Assessment Test.

RELATIONSHIPS AMONG A MEASURE OF AUDITORY AND VISUAL SEGMENTATION AND STANDARDIZED AND INFORMAL MEASURES OF READING READINESS Order No. 8016563

Nolting, L. Wray, Ed.D. *The University of Florida*, 1979. 74pp. Chairperson: Dr. Linda L. Lamme

This study investigated the relationships among scores of sixty-four, randomly selected, kindergarten children, from eight randomly selected classrooms, on the Informal Test of Auditory and Visual Segmentation, with scores on the Metropolitan Readiness Test, reading readiness component, the Basic School Skills Inventory, reading readiness subtest; teacher ratings of reading readiness; and the number of words read correctly in a one-minute sample from an unfamiliar passage. The Informal Test of Auditory and Visual Segmentation is comprised of subtests measuring three metalinguistic abilities: Auditory Word Boundaries, Auditory Syllabication, and Visual Word Boundaries. Multiple regression and Pearson Correlation coefficient techniques were used to analyze the data.

The results indicated that all three metalinguistic abilities assessed in this study, as well as the total Informal Test of Auditory and Visual Segmentation score, are significantly correlated with the standardized and informal measures of reading readiness used in this investigation, the only exception being the ability to determine auditory syllabication which was not significantly correlated with the number of words read correctly in a one-minute sample from an unfamiliar passage.

Sex and age were found to have no effect on the total Informal Test of Auditory and Visual Segmentation score. The results indicated that race may have some effect on the Informal Test of Auditory and Visual Segmentation score.

The reliability of the Informal Test of Auditory and Visual Segmentation was determined to be .81. This investigation also concluded that further research on metalinguistic abilities, their relationships to reading readiness, methods of teaching these skills, and their assessment is essential.

THE DEVELOPMENT AND EVALUATION OF A WORDS-IN-CONTEXT TEST TO DETERMINE INSTRUCTIONAL READING LEVELS

Order No. 8020384

OWEN, JO ANNE, ED.D. *Memphis State University*, 1980. 136pp. Major Professor: Robert A. Kaiser

Purpose. The purpose of this study was to develop a process that would lead to the development of a valid instrument that could be used to determine instructional reading levels for fifth and eighth grade students.

Procedures. This study included a survey of related literature on the necessity of establishing agreement between instructional reading levels and reading materials, the availability of instruments to determine this agreement, and the extent to which these instruments measure instructional reading levels; the procedures and rationale used to select the subjects and to administer and score the instruments; the statistical analysis of the collected data; and conclusions with implications and recommendations for continued research.

The instrument that was developed and evaluated was the *Knowledge of Words-In-Context* test (KWIC). The KWIC is essentially a combination of graded sentences using a sentence completion procedure and high frequency words from a graded word list. It consists of 8 sections referred to by grade levels two through nine with each level consisting of 10 sentences for a total of 80 items.

The sample for this study comprised 370 fifth grade students and 424 eighth grade students. The sample was drawn from four elementary and four junior high schools in a major urban center of the Southeastern United States.

The total sample of 794 was administered the KWIC test, the *Gates-MacGinitie Reading Tests*, and the Cloze Test. The individual *Silvaroli Classroom Reading Inventory* was administered to ten randomly selected students in each school. These 40 fifth grade students and 40 eighth grade students comprised 10% of the total sample.

The assessment of the concurrent validity of the KWIC test was handled by correlating the scores of the subjects on the KWIC test with scores obtained from a group administered standardized reading comprehension test, the *Gates-MacGinitie Reading Test;* a group administered standardized procedure, a cloze test, and for a subsample with scores obtained from an individually administered instrument, the *Silvaroli Classroom Inventory.*

Conclusions. (1) The KWIC can be used for the placement of fifth and eighth grade students in reading groups and for the initial identification of fifth and eighth grade students likely to require more thorough diagnostic inquiry. (2) The construction process applied in this study can lead to the construction of valid and, therefore, useful instruments for the placement of fifth and eighth grade students likely to require more thorough diagnostic inquiry. (3) Race and sex variables make no useful contribution to improving the validity of the KWIC, although race showed significant correlations with standardized diagnostic instruments for the assessment of grade placement reading levels.

Recommendations. (1) It is recommended that tests similar to the KWIC may be constructed and evaluated for other grades by using the process developed in this study. (2) It is recommended that tests using the KWIC process may be developed in the content areas and evaluated.

THE RELATIONSHIPS BETWEEN SELECTED DISCOURSE ANALYSIS TASKS, PREVIEWING STRATEGIES AND SUBSEQUENT FIFTH AND SIXTH GRADE STUDENTS' PERFORMANCE ON MEASURES OF READING COMPREHENSION

Order No. 8019562

RAUCH, MARGARET MARIE, PH.D. *University of Minnesota*, 1980. 205pp.

The major purpose of this study was to investigate the effectiveness of two instructional methods on the performance of better and poorer readers on measures of reading comprehension. Results from research indicate that prior knowledge positively influences comprehension. Thus, one method, previewing, was based upon the premise that helping students integrate prior knowledge with the topic to be read would facilitate comprehension. There is also evidence from research that text structure has an effect on comprehension. The second method, discourse analysis, was based upon the premise that helping students attend to the relationships between and within propositions in the microstructure of text would facilitate comprehension.

There were two experiments in this study. Experiment one involved sixty-three fifth grade students from a private elementary school. Experiment two was a replication of experiment one and involved twenty sixth grade students. Subjects were blocked on reading achievement (high, low) based on their pretest total scores from a standardized reading test and then randomly assigned to one of two instructional methods. Instruction was scheduled for one-half hour time periods and extended for twenty days.

Eighteen expository selections were used during treatment. Multiple choice questions were generated on the detailed propositions of each selection. Two question types were used. Implicit question types were based on propositions represented in the surface structure of the text. Implied question types were based on propositions which were not represented in the surface structure. In particular, the questions focused on four different connective propositions (causality, purpose, time, location).

Six dependent measures were used: the vocabulary and comprehension components from a standardized reading test, tests on main ideas and details, and the implicit and implied components from a test of propositional analysis. Results from the multivariate analyses of variance for both experiments showed no significant difference between instructional methods. Several factors may have contributed to this outcome. Treatment may not have been of sufficient duration to effect changes on standardized reading measures. Selections used during treatment were judged to be within the reading level of the students, thus, students may not have been expanding their vocabularies to the extent needed to effect changes on a standardized reading measure. Four types of connective propositions were focused on in the discourse method and instruction may be needed in either all or different types of connective propositions.

Results from the paired t-tests for both experiments revealed that both better and poorer readers achieved a significantly higher performance on implicit than on implied question types.

Results from the multivariate analyses showed a significant difference between better and poorer readers in experiment one but not in experiment two. A two-way analysis of variance showed a significant difference between better and poorer readers in experiment two at the beginning of treatment. An inspection of the univariate analyses revealed that the nonsignificant difference obtained on the implicit component from the test of propositional analysis contributed the most toward the result obtained on the multivariate analysis. There was one interaction in experiment two. On the test focusing on the main idea, the better and poorer readers in the previewing method performed in the same manner. In contrast, the better readers in the discourse method achieved the highest performance and the poorer readers had the lowest performance. One possible explanation may be that the previewing method helped the poorer readers and the discourse method did not help the poorer readers to the same degree.

Classroom implications and further needs for research were discussed.

AN EXPERIMENT FOR DETERMINING THE MAIN IDEA IN SECONDARY SCHOOL READING: THE SUBSKILL METHOD VERSUS THE MODIFIED LANGUAGE EXPERIENCE METHOD

Order No. 8020678

ROWE, M. ELISE MURPHY, ED.D. *Northern Illinois University*, 1980. 175pp.

The major purpose of this study was to examine the effectiveness of two methods for determining the main idea in secondary school reading. Approaches compared were The Subskill Method and The Modified Language Experience Method.

Concentration in The Subskill Method centered on moving from the specific understanding of the main idea of the paragraph to the general comprehension of the paragraph. The Modified Language Experience Method stressed progressing from the general comprehension of the paragraph to the specific understanding of the main idea of the paragraph. Each paragraph used in The Subskill Method was prepared professionally. Paragraphs concerned with The Modified Language Experience Method came from two sources. The sources were language experience stories written by the subjects themselves and paragraphs prepared professionally.

Pupils taking part in the study were freshman students in a Chicago suburban area public high school. All subjects took three tests.

The first test was the Nelson-Denny Reading Test, Form C, which was used to select individuals whose reading comprehension was on or above ninth grade level. The second test was the Six-Way Paragraph Criterion Reference Test, Form I, Main Idea, which was used both as a pretest and a posttest. As a pretest it was given to students reading on or above ninth grade level to determine which pupils scored lower than 80 percent on determining the main idea. As a posttest it was used to determine if there was a significant difference in results for the following two groups in determining the main idea: The Subskill Group and The Modified Language Experience Group.

Data provided from the Nelson-Denny Test and the Criterion Reference Pretest were ranked to provide an even distribution of qualifying subjects. Subjects were divided into two groups. The odd numbered group received The Subskill Method of teaching the main idea while the even numbered group received The Modified Language Experience Method of teaching the main idea.

In order to acquire further information and to determine if a second study would result in the same or a different conclusion, two separate studies were conducted. A different population from the same school was used in each study.

The Original Study took place the Spring semester of 1979, while the Replication Study was administered during the Fall semester of 1979. Each semester approximately 300 students were examined. After testing, the qualifying target population for the Original Study was 64 students while the qualifying target population for the Replication Study was 44 students.

Statistical analysis led to the conclusion that there was not a significant difference between the two methods of teaching the main idea. Students receiving The Subskill Method for determining the main idea and students receiving The Modified Language Experience Method for determining the main idea made almost similar growth. The similarity was present in both the Original Study and in the Replication Study. Research indicates, therefore, that both The Subskill Method and The Modified Language Experience Method can be productive in high school reading comprehension.

Recommendations of the study suggest continued research in the methods of determining the main idea. Specific studies could concentrate on additional ways The Subskill Method and The Modified Language Experience Method could be applied to teaching students to determine the main idea in content area reading.

A CANONICAL CORRELATIONAL ANALYSIS OF THE READING-WRITING RELATIONSHIP: AN EXPLORATORY INVESTIGATION

Order No. 8019947

SHANAHAN, TIMOTHY EDWARD, PH.D. *University of Delaware*, 1980. 171pp. Professor in charge: John J. Pikulski

Examines the relationship of learning to read and learning to write at the second and fifth grade levels. Measures of phonics knowledge, reading vocabulary, spelling, reading comprehension, and grammatical and organizational complexity of writing were administered to 256 second graders and 253 fifth graders. Tests were selected on the basis of their ability to provide a reliable measurement of important reading and writing skills which have been found, in previous bivariate studies, to be interrelated. Grades two and five were selected to maximize differences in the relationships which might be due to development or learning. Data was analyzed through a canonical correlational factor analysis.

Major findings were: (1) A single significant canonical factor best summarized the reading-writing relationships at both grade levels; (2) These canonical factors were able to account for significant proportions of reading and writing variance (25-40%); (3) The nature of the reading-writing relationship, as described by these factors, is different at the second and fifth grade levels. At both grade levels, the spelling and word recognition variables unite in a single dimension. However, this relationship is based on both visual and phonemic aspects of spelling ability at the second grade level; but it is only a visual spelling factor at the fifth grade level; (4) Reading level differences distinguish more clearly the nature of the reading-writing relationships than do the grade level differences. This study suggests that the reading-writing relationship for children reading below the third grade level is best described as a word recognition-word production (spelling) relationship. For proficient readers, fifth grade level and above, the relationship is more a reading comprehension-prose production relationship. The relationship at this level is based largely on the vocabulary diversity and organizational complexity of children's writing.

THE DEVELOPMENT AND ANALYSIS OF A CRITERION-REFERENCED TEST OF TEXTBOOK READING AND VOCABULARY DISCRIMINATION FOR UNIVERSITY FRESHMEN WITH REFERENCE TO THEIR ACT COMPOSITE SCORES

Order No. 8019380

SMITH, KAREN GAIL, ED.D. *New Mexico State University*, 1980. 160pp. Chairman: Dr. Leon E. Williamson

The purpose of this study was to develop a criterion-referenced test of textbook reading and vocabulary discrimination for university freshmen.

Achievement objectives for four domains in vocabulary discrimination and two in skills of using textbooks were written as a base upon which the test could be designed: acceptable letter groups, syllabication, phonological spelling, use of word parts, use of textbook parts, and comprehension of textbook selections. A rationale was presented for each of the subtests of the instrument. The domain of each subtest was described, and a pool of items was designed for each.

A pilot study was conducted in order to evaluate content validity. The instrument was refined using the information generated by the pilot study. Consensus guidelines for evaluating the performance on the textbook comprehension subtest were developed with the assistance of a panel of judges. A competency of 80 percent accuracy on each of the subtests was set, and a confidence interval of 10 percent was accepted which generated a minimum criterion standard of 70 percent. Those achieving 70 percent were defined as masters.

The sample selected for this study was a group of 40 first-year students at New Mexico State University. Twenty students each were randomly selected from the bottom and top thirds of the ACT composite scores so that the instrument could be administered to two distinctly different groups.

Each subtest of the instrument was scored and investigated independently of the other subtests. Comparisons were then made between group performances on each part, and the parts were evaluated in relation to each other. Subtests were investigated using raw scores, percentages, and means. Items within subtests were investigated with the use of an analysis of variance test and a descriptive analysis procedure which analyzed accuracy, content, and efficiency. Those items which discriminated between the two groups at the 5 percent level of significance were automatically accepted as accurate, content-specific, and efficient. However, the traditional determination of significant response variance is not required in criterion-referenced test development; thus, some reduced variance was accepted with items otherwise analyzed as accurately written, content-specific, and somewhat efficient.

Within the limitations of the study, the instrument identified masters and nonmasters of two different groups of subjects. The high ACT group performed better, to varying degrees, on each of the subtests. The low ACT group were identified as masters on 24 percent and the high group on 57 percent of the subtests. A bimodal distribution was demonstrated in frequency of mastery of the subtests. Domain-selection validity was demonstrated in that subtest IV, which measured knowledge of word parts, was 100 percent generalizable to four of the subtests and 66 percent generalizable to the fifth subtest.

Potential exists for the criterion-referenced test developed in this study. Recommendations are: (1) That a revision be made of those faculty items identified through analysis of variance and the descriptive analysis; (2) That a further analysis be done on the textbook selections used in measuring textbook reading comprehension; (3) That an investigation of the writing samples produced in responses to the section on textbook comprehension be initiated; (4) That for future samples, information on the influence of a first language other than English be included; (5) That after another refinement of the instrument, analysis of relationships possibly existing between subtests be investigated; (6) That a follow-up study on the academic success of the original 40 subjects be conducted to determine criterion-referenced validity; (7) That functional validity of the instrument be determined after the test has been used as a preassessment and postassessment device.

AN INVESTIGATION OF THE AURAL CLOZE PROCEDURE AS A MEASURE OF LISTENING COMPREHENSION

Order No. 8018467

SMITH, VIRGIL DEMPSY, II, PH.D. *Arizona State University*, 1980. 147pp.

Listening has been the concern of educators for decades. Due to the long and cumbersome nature of traditional listening tests, some individuals have suggested that an investigation of the aural cloze procedure as a measure of listening comprehension might prove beneficial. To these ends a study was devised which involved 240 ninth graders who listened to or read cloze passages from the content fields of social studies and science with readability of fifth, seventh or ninth grade levels. The mixed cross-nested design yielded results that indicated: (1) students were more successful when responding to the written cloze procedure, (2) the aural and written cloze procedures were affected in a similar manner by the external variables of readability and content, (3) the student-related variable of reading score affected both cloze performances, (4) the aural and written cloze testing procedures were reliable and equitable measures of comprehension abilities, and (5) the aural cloze procedure was a reliable measure of aural language comprehension abilities in this study.

THE RELATIONSHIP BETWEEN DEMONSTRATED STORY GRAMMAR USAGE BY THIRD GRADERS AND THEIR SCORES ON SELECTED READING COMPREHENSION TESTS

Order No. 8024162

SUMMERS, PAMELA FRENCH, ED.D. *Boston University School of Education*, 1980. 165pp. Major Professor: J. Richard Chambers

The three major purposes of this study were to investigate and ascertain if: (1) there is a relationship between story generation and scores received on standardized reading comprehension tests; (2) there is a relationship between story recall and scores received on standardized reading comprehension tests; and (3) there is a relationship between story grammar usage (story generation and recall) and reading achievement.

Subjects for the study were thirty third grade students from a suburb of Burlington, Vermont. Based on the *Stanford Achievement Test*, the subjects were divided into groups reading above, on and below grade level.

From school records, scores on the *Stanford Achievement Test* Total Reading, Word Reading plus Reading Comprehension, and Reading Comprehension sub-tests were obtained, as well as the Independent Reading Level achieved on the *Sucher-Allred Reading Placement Inventory*. The *Durrell Listening-Reading Series* test was administered at the time of the study.

Story Grammar usage was measured by two instruments. Each child was asked to tell a good story from a single story stem, which was taped and transcribed. Each protocol was rated and assigned a story generation score using a scoring guide developed for the study. Story recall assessment was derived from each child's recall of narratives written according to the Stein and Glenn Story Grammar Analysis and rewritten with varying modifications (deletions and reordering). Recall protocols were scored according to the number of propositions recalled and the completeness of recall (measured by ordering and/or insertions of category information).

Pearson's product moment correlations and analyses of variance were performed to determine the relationships between standardized reading comprehension tests and story grammar usage.

A content analysis of the passages designed to test reading comprehension on the standardized tests utilized in the study was performed, applying the same scoring guide designed to rate the generated stories.

Findings. That data showed that the relationship between stories generated and recall of narratives by third grade students and the standardized reading comprehension tests is positive but not significantly so. A low, but negative correlation was found between age and story grammar within a single grade level. The differences between reading achievement and story generation were significant. However, when pre-existing differences in age were controlled by analysis of covariance, the difference was not significant. The difference between the recall of stories by reading achievement levels was not significant. A content analysis showed that only 22% of the total number of passages designed to test reading comprehension on the standardized tests could be measured as story structures.

Conclusions. Based on the findings, it was concluded that reading passages designed to measure reading comprehension on standardized tests do not contain large percentages of well formed stories; there is a low but positive correlation between story grammar usage and standardized reading comprehension test scores; younger children who are reading above grade level generated higher order story structures than older children reading on or below grade level; reading achievement has a low but positive affect on story generation; and reading achievement does not significantly affect the amount of story recall.

ORAL READING ERRORS OF FIFTH GRADE STUDENTS

Order No. 8014568

TURNER, GLADYS H., ED.D. *Temple University*, 1980. 98pp.

Purpose. This investigation was conducted to determine if there were differences among the mean percentages of graphic, syntactic, and semantic errors and graphic, syntactic, and semantic self-corrections made by groupsof High W. K. subjects and groups of High Camp subjects in the fifth grade as they read orally.

Procedure. Of the 225 fifth grade students in a suburban-rural school, 49 obtained scores which were eight or more months higher on the work knowledge subtests than on the comprehension subtest of the Metropolitan Achievement Test (MAT). They were known as High W. K.'s. The 44 subjects who scored eight or more months higher on the comprehension subtest than on the word knowledge on the comprehension subtest than on the word knowledge subtest were known as High Comp's.

The 93 subjects were further divided into one of three groups dependent on their total reading scores as measured by the stanines obtained on the MAT. Group I contained the subjects who obtained stanines of 7, 8, orr 9. Group II contained the subject who obtained stanines of 4, 5, or 6. Group III contained the subjects who obtained stanines of 1, 2, or 3. Only five High W. K.'s were in this Group.

During the last three weeks of the school term, the students individually read orally-at-sight without any help a 455 word article, "Reading, What Is It?" from a fifth grade basal reader. The tape recorder was activated and the examiner recorded the errors on xeroxed copies of the story. At the conclusion of all the testing, the tapes were replayed as many times as was necessary in order to check the record of the errors and the self-corrections on the xeroxed copies.

All of the tyypes of errors were totaled and divided by the total number of errors of each subject. The same computation was done to find the percentages for types of self-corrected errors except the divisor was the total of self-corrected errors.

A univariate t-test was performed for each dependent variable for each group. Hotelling's T^2, a multivariate test, was performed for the total High W. K.'s and the total High Comp's.

Findings. The High W. K.'s in Group I made a significantly higher mean percentage of syntactic errors than the High Comp's. This was not true for Group II.

Hotelling's T^2 showed that the three dependent variables of self-corrected errors differentiated the High W. K. subjects and the High Comp subjects at an alpha level approaching significance ($T^2 = 8.04$, $F = 2.62$, $df = 3/89$, sig. $= .055$).

The High W. K.'s made a higher percentage ($p < .01$) of graphic self-corrections and a higher mean percentage of syntactic self-corrections ($p < .05$) than the High Comp's.

As the 93 subjects read the 455 words, they miscalled 250 words, omitted 138 words, added 169 words, and reversed three phrases. The mean number of errors was 28.6 with a range of 9 to 87 based upon a total of 2665 errors. The mean number of self-corrected errors was 3.34 with a range of 0 to 11 based upon a total of 311 self-corrected errors.

Conclusions. The total High W. K. subjects made a significantly higher mean pecentage of graphuic and syntactic self-corrections than the total High Comp subjects. Therefore, it appeared that the subjects who obtained scores of eight months or more higher on the MAT subtest of work knowledge than on the subtest of comprehension used different strategies for processing information while reading orally than the subjects who obtained scores eight months or more higher on the MAT subtest of comprehension than the subtest of word knowledge.

THE EFFECTS OF DIFFERING RESPONSE CRITERIA ON THE ASSESSMENT OF WRITING COMPETENCE Order No. 8016064

WINTERS, LYNN, ED.D. *University of California, Los Angeles*, 1979. 174pp. Chairperson: Professor Eva L. Baker

The method selection for judging writing samples *ipso facto* defines "good writing." Decisions, however, about writing competence derived from one scoring system may or may not lead to the same decisions about an examinee when another system is used. The effect of competing methods of judging writing samples on the classification of students into appropriate levels of writing skill is examined with four research questions: (1) What is the pattern of performance of different kinds of writers on different scoring systems? (2) What is the pattern of performance of these writers on the same system? (3) What scoring system best discriminates among writing groups? (4) Which scoring systems are the best predictors of criterion group status?

Four contrasting scoring systems, T-unit analysis, the Diederich Expository Scale (DES), the CSE Analytic Scale (CSE) and General Impression marking (GI) were selected to represent frequency counting, analytic, and holistic approaches to essay scoring. Four criterion groups of twenty students each from high school and college composition classes served as a priori classification groups against which the four scoring criteria were tested.

The eighty students in the study produced two 200-word expository writing samples from parallel topics on two test occasions, one week apart. Topics and test occasions were counterbalanced within criterion groups. Essays were read in four rating sessions, one for each scoring system, by trained raters. All essays received two ratings. The four dependent measures consisted of each student's total score (the sum of two ratings across two topics) for each scoring system.

Several comparisons were made among the systems. Generalizability coefficients were highest for DES, .85, comparable for GI and CSE, .63 and .67 respectively, and undefined for T-unit, due to large subject by topic effects. The shapes of the score distributions produced by each system differed within and between groups. Group means on each system ranked from high to low as follows: College Low, College High, High School High and High School Low. Combined scores from GI, DES, and CSE provided the best overall classification accuracy, 61%. A combination of DES-CSE scores produced an accuracy of 59%; accuracies of the individual systems were 53% for DES, 51% for CSE, 46% for GI and 24% for T-unit.

When classification accuracies for individual criterion groups were examined, the GI-DES-CSE combination of scores was best for predicting membership in all groups but the College High. For this group, DES alone was the best predictor. Interestingly, the group by system interactions found in the classification analyses were evident in the intercorrelations among system. Correlations among the systems differed for each of the criterion groups suggesting the level of writing ability interacts with the kind of criteria selected to judge essays.

Results of this study indicate that good writing is not defined in the same manner for different levels of writers. Further research is needed to isolate those scoring systems that best describe and classify high school and college writers. Until the classification validities of several sorts of criteria are established, writing assessment programs designed to make diagnostic-placement decisions about students with different levels of writing skill, will lack credibility and integrity.

Copies of the dissertations may be obtained by addressing your request to:-

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106

or by telephoning (toll-free) 1-800-521-3042